Animating Pictures of Cooking using Video Example of Bubble and Image Deformation

Sakiyama, Syohei
The University of Electro-Communications

出版情報：MI lecture note series. 50, pp.74-77, 2013-10-21. 九州大学マス・フォア・インダストリ研究所
バージョン：
権利関係：

Animating Pictures of Cooking using Video Example of Bubble and Image Deformation

Syohei Sakiyama
The University of Electro-Communications, Tokyo, Japan
(joint work with Makoto Okabe and Rikio Onai)

1. Introduction

We see pictures of cooking or dishes in the web site of a restaurant. Fig. 1-left shows the example of a picture of hot pot. We want to replace such pictures with the videos, since the video allows understanding more efficiently how the hot pot is delicious by showing rising steam, bubbles, and sizzling sounds.

However, it is not a good idea to directly record the video of cooking a hot pot. Fig. 1-right shows the video of a boiling hot pot, but it does not look delicious compared with the picture: the chickens and vegetables have changed their shapes and sunk down. Also, the overall appearance has lost its original vividness.

To create the delicious-looking animation of cooking and dishes, we propose an interactive method to animate the picture like Fig. 1-left. The user begins to animate a picture by interactively specifying the regions of bubbles and the other ingredients, and adjusts several parameters for each region. The system then synthesizes the animation by making the composite of the input image and the video examples of bubble. We demonstrate that our system enables users to create the comparably high quality animation in less time compared with the standard video editing software.

Figure 1. The left is the picture of a hot pot. The right is the video of an actual cooking hot pot.

2. System Overview

Given the target image of cooking or dishes (Fig. 2-a), the user begins to create the animation by specifying the regions of bubbles and foods. We use a sketch-based image

segmentation tool [1]: drawing with the left mouse button specifies the bubble regions and the right mouse button specifies the foods (Fig. 2-b). Using the same tool, the user further separates the food region into detailed ingredients like meats and vegetables (Fig. 2-c). For each ingredient region, the user selects one of the two parameters; *deformation pattern* and *shape type* (Fig. 2-d). For the deformation pattern, the user selects *deformable* or *non-deformable* pattern: deformable region is suitable to soft ingredients like sliced meat or green leaves that will deform a lot. Non-deformable region is suitable for hard ingredients like diced chicken. For the shape type, the user selects one of five types of *plane*, *sphere*, *stick*, *leaf*, or *noodle*. Given these information, our system automatically synthesizes the animation by making the composite of videos examples of bubble. The parameters such as the position of a bubble, its size, its speed, and the vibration speed of an ingredient are automatically chosen. The user can fine-tune the parameters, if necessary.

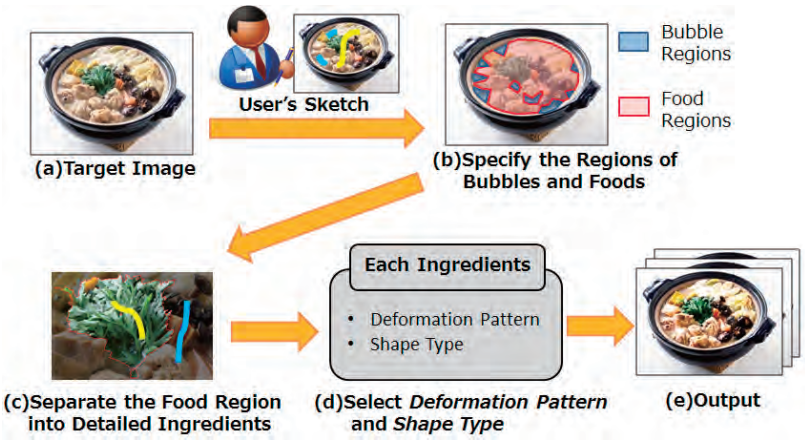

FIGURE 2. System overview.

3. PLACEMENT OF VIDEO EXAMPLES OF BUBBLE AND ANIMATION SYNTHESIS

Since the bubbles are important to animate the picture of cooking, we prepare the database of video examples of bubble. We record the several types of real bubbles using a video camera by ourselves. Fig. 3) shows three examples of our bubble videos. We also edit each video to prepare its multiple versions by changing the scale and playback speed.

Our system automatically selects suitable bubbles for each bubble region in the input image. For this purpose, we observe the actual bubbles during the cooking of a hot pot, and contrive our algorithm. We place the video examples of bubble according to the rules of Table 1).

FIGURE 3. Our database includes the videos of boiled (a) water, (b) ketchup, and (c) soy sauce.

TABLE 1. The algorithm for placement of bubbles

Step1	A few large high bubbles at the center of the hot pot
	A few small high bubbles in the large region far from the center
	Many small high bubbles in the small region far from the center
Step2	High bubbles along the boundaries of the food region
	High bubbles in the narrow region between ingredients

We synthesize the animation by making the composite of the input image and the video examples of bubble. The synthesis is done in HSV color space: we use the H and S channels of the input image and the V channel of the video example to synthesize the animation. We adjust the V channel of the video example so that the brightness of it matches to the brightness of the input image. The process is shown in Fig. 4. For each pixel position (x, y) in t-th frame of the video example, we let $V_t^I(x, y)$ the brightness of the input image and $V_t^V(x, y)$ the brightness of the video example. We compute the brightness of the synthesized animation $V_t^A(x, y)$ as follows:

$$V_t^A(x,y) = V_t^I(x,y) + V_t^V(x,y) - V_{avg}^V(x,y), \tag{1}$$

$$V_{avg}^V(x,y) = \frac{\sum_{t=1}^{F}\sum_{x,y} V_t^V(x,y)}{N \cdot F}, \tag{2}$$

where N is the number of pixels and F is the number of the frames of the video example. This method enables the natural animation synthesis, even if the overall appearnce of the video example is different from the input image.

4. VIBRATION OF INGREDIENTS BY RIGID TRANSFORMATION

In the hot pot, each ingredient is vibrated by the rising bubbles. Our system synthesizes such vibrating animations based on a few user-specified parameters. The user's inputs are the deformation pattern and shape type described in Fig. 2-d. The ingredient of "deformable" pattern vibrates by deforming the shape. The ingredient of "non-deformable" pattern vibrates by not deformation but only translation.

We use the rigid transformation using moving least squares method [2] for the deformation. We place several control points and deform the image by moving the control points. We place control points according to the shape type. Fig. 5 shows

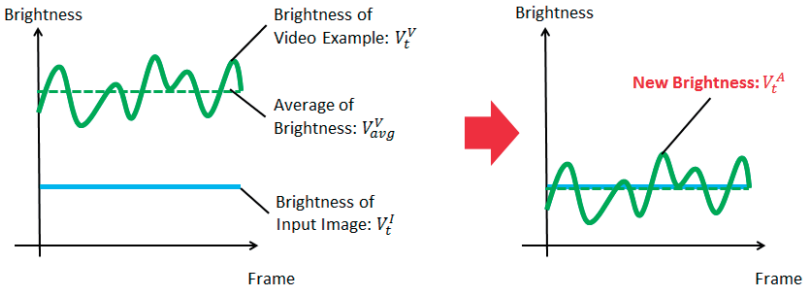

FIGURE 4. Adjustment of the brightness in V channel.

examples. Given the image of a sliced meat of *plane* type, the system places six control points. Four of them are always fixed (as shown in red color) and two of them vibrate (as shown in yellow color).

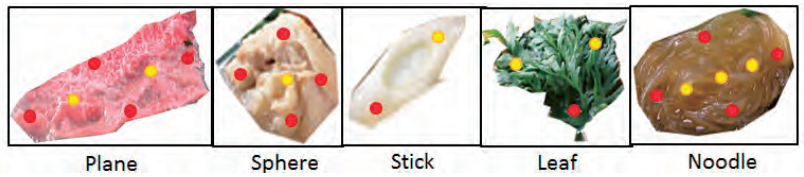

FIGURE 5. Placement of control points for each shape type. Red points are fixed, and yellow points vibrate.

5. RESULTS

We use our system and Adobe After Effects (AAE) to synthesize the animation of the picture of the hot pot (Fig. 2). To synthesize the animation of the hot pot, the user spends ten minutes using our system, and six hours using AAE: our system is significantly superior to the standard video editing software from the viewpoint of the usability. From the viewpoint of the quality of the resulting animation, our result is slightly more unnatural than the result of AAE. Especially, the vibration in our result looks worse. We want to improve it by designing the vibration algorithm more carefully in the future.

REFERENCES

[1] T. Kiyono, T. Hayashi, R. Onai, M. Sanjo, and M. Mori: Proposal and Evaluation of Fast Image Cutout Based on Improved Seeded Region Growing, *Information Processing Society of Japan*, Vol.50, No.12, 3233-3249, 2009.

[2] S. Schaefer, T. McPhail, and J. Warren: Image Deformation Using Moving Least Squares, *In Proceedings of SIGGRAPH 2006*, 533-540, 2006.